THE TWO TREATMENT TWO PERIOD CROSSOVER DESIGN: A REVISITATION USING A FULL RANK CELL MEANS MODEL

BU-669-M March 1979

R. P. Kershner

Abstract

A cell means formulation of the general linear model under a correlated error structure is introduced together with some of its properties relating to estimation of population parameters. The utility of this formulation is illustrated by reworking the results in Grizzle (1965) regarding estimation and testing in a two treatment two period crossover design when residual effects are present.

1. Introduction

The two treatment two period design enjoys extensive use in applications. It is particularly popular in the pharmaceutical industry where a test drug is to be compared with a standard on a cohort of human subjects. An advantage of the crossover design is that of having comparisons between treatments made on a within subject basis. Each subject in effect serves as his own control, and, under certain conditions, crossover designs can be more efficient than their completely randomized counterparts.

Whenever sampling units are used repeatedly under a sequence of one or more treatments there exists the possibility that treatment effects may last beyond the period of application. In the literature these effects are called carry over or residual effects. There can exist $1^{st}, 2^{nd}, \cdots, k^{th}$ order residual effects which last respectively $1, 2, \cdots, k$ periods beyond the actual treatment period. Direct treatment effects correspond to treatment effects which are observed only in the period of application.

The literature on crossover designs with resiudal effects is concerned primarily with construction of classes of designs having certain optimal properties with respect to estimation of direct and 1^{st} order residual effects. Examination of classes, or members of classes, of two treatment crossover designs has been largely neglected.

The work of Grizzle (1965) is significant in the sense that it is the only paper to closely examine the two treatment two period crossover design when residual effects are present. He showed that residual effects are not estimable in the presence of period effects.

Brown (1978), as well as unpublished reports coming out of the Food and Drug Administration, caution against using the two treatment two period design when residual effects are present. There appears to be no generalized treatment of possible solutions to the problem of estimability in the two period crossover design providing both efficient alternatives to the design and giving a unified treatment of their analysis in more complex settings.

As a first step toward developing such a unified approach to the problem, this paper introduces a general full rank linear model applicable to all crossover designs. The basic structure and properties of this formulation are made in a completely general setting; i.e., p periods, s sequences, n_i subjects in sequence i, i=1,$\cdots$,s, and where the observations on a given individual have arbitrary covariance matrix $\underset{\sim}{\Omega}$. In this setting the model is in fact a multivariate cell means model. The direct applicability of this model to crossover designs is shown in this paper by reworking the results in Grizzle (1965). Specific generalizations to more complex crossover designs ($t > 2$, $p > 2$) will be developed in a subsequent paper.

2. The Design and Model

The two treatment two period crossover design is simply a latin square of order two where the rows correspond to periods and the columns to sequences of treatment administration AB and BA where the two treatments are designated as A and B. Sampling units (s.u.'s, e.g., humans or animals) are randomly assigned to sequences in such a way that half the s.u.'s receive the AB sequence and half the s.u.'s receive the BA sequence. Over the course of the experiment, each s.u. will receive both treatments. Schematically the design is as follows:

	Sequence	
Period	1	2
1	A	B
2	B	A

A model for this design might be one which is simply an extension of the conventional model for a latin square with the addition of residual effects. Thus it would be

$$y_{ijk} = \mu + \pi_i + \beta_j + \tau_\alpha + \delta_{2i}\rho_j + \xi_{k(j)} + \epsilon_{ijk} \quad , \tag{2.1}$$

where i=1,2, j=1,2, k=1,$\cdots$,n_j and, for y_{ijk} being the observation corresponding to the k^{th} subject at period i in sequence j,

μ: an overall mean,

π_i: effect due to period i,

β_j: effect due to sequence j,

τ_α: effect due to treatment α where

$$\alpha = \begin{cases} 1, & \text{for treatment A,} \quad \text{where } i = j \\ 2, & \text{for treatment B,} \quad \text{where } i \neq j \end{cases}$$

ρ_j: residual effect due to j^{th} treatment and δ_{2i} denotes the usual Kronecker delta such that

$$\delta_{2i} = \begin{cases} 0 & i=1 \\ 1 & i=2 \end{cases},$$

$\xi_{(j)k}$: effect due to s.u. k which is nested within sequence j, and

ϵ_{ijk}: random error term corresponding to the ijk^{th} observation.

The colon notation above is to be interpreted as meaning "is defined as".

Except for the addition of sequence effects, (2.1) is the first model given in Grizzle's (1965) paper. It should be noted that his parameterization of (2.1) is incorrect since he has residual and direct treatment effects sharing the same subscript. In addition, one needs the indicator variable δ_{2i} multiplying the residual effects since they do not appear in period 1.

In the parameterization in (2.1), several features are immediately obvious: (a) the treatment effect is completely confounded (or aliased) with the interaction effects of periods and sequences, (b) the residual effects are completely confounded with sequences, and (c) the residual effects are present only in period 2. The net result of all this confounding is that neither $\rho_1-\rho_2$, $\pi_1-\pi_2$, nor $\beta_1-\beta_2$ are estimable functions in (2.1). Grizzle (1965), in reformulating (2.1), explicitly assumes $\pi_1=\pi_2$ and tactily assumes $\beta_1=\beta_2$, thereby giving

$$y_{ijk} = \mu' + \delta_{2i}\rho_j + \tau_\alpha + \xi_{k(j)} + \epsilon_{ijk} \qquad (2.2)$$

where, for the ij^{th} cell, $\mu' = \mu + \pi_i + \beta_j$.

The usual distributional assumptions made about the elements of (2.2) are that

$$\xi_{(j)k} \text{ are i.i.d. } n(0,\sigma_s^2)$$

and, independently, by

$$\epsilon_{ijk} \quad \text{are i.i.d.} \quad n(0,\sigma_e^2) \quad .$$

On defining $\sigma^2 = \sigma_e^2 + \sigma_s^2$, the covariance structure on the y_{ijk}'s is then

$$\text{cov}(Y_{ijk},Y_{i'j'k'}) = \begin{cases} \sigma^2 & i=i',\ j=j',\ k=k' \\ \sigma_s^2 = \rho\sigma^2 & i=i',\ k=k',\ j\neq j' \\ 0 & \text{otherwise} \end{cases} \quad , \qquad (2.3)$$

and where

$$\rho = \text{corr}(Y_{ilk},Y_{i2k}) \quad .$$

Without loss of generality, take $\underset{\sim}{y}'$ to be the 1 x 2n vector of observed values in lexicon order where $n = n_1 + n_2$. The covariance structure is then

$$\underset{\sim}{\Sigma} \equiv \text{var}(\underset{\sim}{y}) = \underset{\sim}{\Omega} \otimes \underset{\sim}{I}_n$$

where $\underset{\sim}{\Omega} = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ and $\otimes$ denotes the usual right Kronecker product such that $A \otimes B = \{a_{ij}B\}$ [e.g., Searle, (1966), p. 215].

3. The Cell Means Model

In a completely general setting a crossover design will have p periods, s sequences, and n_j s.u.'s in the j^{th} sequence for j=1,···,s. Denote the total number of distinct s.u.'s vs $n = \Sigma_{j=1}^{s} n_j$. The generalized cell means model [e.g., Hocking and Speed, (1975); Speed and Hocking, (1976); Speed, Hocking and Hackney, (1978)] is then $Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$ where μ_{ij} is the population mean corresponding to period (row) i and sequence (column) j. In matrix form the model is

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\mu} + \underset{\sim}{\epsilon} \tag{3.1}$$

where $\underset{\sim}{y}'$ is the 1 × pn vector of observed values in lexicon order

$$X = I_p \otimes \left(\overset{s}{\underset{j=1}{\Sigma^+}} \underset{\sim}{1}_j \right) \tag{3.2}$$

and $\underset{\sim}{\mu} = \{\mu_{ij}\}$ is the vector of population means in lexicon order. In (3.2) the notation Σ^+ is a generalization of a direct sum [e.g., Searle, (1971), p. 231], and $\underset{\sim}{1}_j$ denotes an n_j × 1 vector containing all 1's.

On generalizing the covariance assumptions in (2.3) to p observations per subject, it follows that:

$$\underset{\sim}{\Sigma} : \text{var}(\underset{\sim}{\epsilon}) = \underset{\sim}{\Omega}_{p\times p} \otimes \underset{\sim}{I}_{sn} \quad , \tag{3.3}$$

where

$$\underset{\sim}{\Omega} = \sigma^2\{(1-\rho)\underset{\sim}{I} + \rho\underset{\sim}{J}\} \quad .$$

Theorem: For the model $y = X\beta + \epsilon$ with

$$\text{cov}(\underset{\sim}{\epsilon},\underset{\sim}{\epsilon}') = \underset{\sim}{V} \quad , \tag{3.4}$$

$$(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y}$$

iff there exists a nonsingular $\underset{\sim}{F}$

$$\underset{\sim}{V}\underset{\sim}{X} = \underset{\sim}{X}\underset{\sim}{F} \quad . \tag{3.5}$$

The proof is straight forward and is given in Graybill. That $\underset{\sim}{\Sigma}$ in (3.3) is a member of this class of dispersion matrices satisfying (3.5) may be seen by taking $\underset{\sim}{F} = \underset{\sim}{\Omega} \otimes \underset{\sim}{I}_s$ since

$$\begin{aligned}
\underset{\sim}{\Sigma}\underset{\sim}{X} &= (\underset{\sim}{\Omega} \otimes \overset{s}{\underset{j=1}{\Sigma^+}} \underset{\sim}{I}_j)(\underset{\sim}{I}_p \otimes \overset{s}{\underset{j=1}{\Sigma^+}} \underset{\sim}{1}_j) \\
&= (\underset{\sim}{\Omega} \otimes \overset{s}{\underset{j=1}{\Sigma^+}} \underset{\sim}{1}_j) \\
&= (\underset{\sim}{I}_p \otimes \overset{s}{\underset{j=1}{\Sigma^+}} \underset{\sim}{1}_j)(\underset{\sim}{\Omega} \otimes \underset{\sim}{I}_s) \\
&= \underset{\sim}{X}\underset{\sim}{F} \quad . \qquad \text{Q.E.D.}
\end{aligned}$$

Here $\underset{\sim}{I}_j$ denotes an identity matrix of order n_j. Henceforth, the limits on the Krunecker sum will be suppressed with the understanding that, unless otherwise noted, the limits are for $j=1,\cdots,s$.

More generally, this correspondence between GLS and OLS can be shown directly for any cell means model defined by (3.2) and (3.3).

Under GLS, using (3.2) and (3.3),

$$\begin{aligned}
\hat{\underset{\sim}{\mu}} &= (\underset{\sim}{X}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \\
&= [(\underset{\sim}{I}_p \otimes \Sigma^+ \underset{\sim}{1}_j')(\underset{\sim}{\Omega}^{-1} \otimes \Sigma^+ \underset{\sim}{I}_j)(\underset{\sim}{I}_p \otimes \Sigma^+ \underset{\sim}{1}_j)]^{-1}(\underset{\sim}{I}_p \otimes \Sigma^+ \underset{\sim}{1}_j')(\underset{\sim}{\Omega}^{-1} \otimes \Sigma^+ \underset{\sim}{I}_j)\underset{\sim}{y} \\
&= [\underset{\sim}{\Omega}^{-1} \otimes \Sigma^+ \underset{\sim}{n}_j]^{-1}(\underset{\sim}{\Omega}^{-1} \otimes \Sigma^+ \underset{\sim}{1}_j')\underset{\sim}{y}
\end{aligned}$$

$$= (\underset{\sim}{\Omega} \otimes \Sigma^{+}\ 1/n_j)(\underset{\sim}{\Omega}^{-1} \otimes \Sigma^{+}\ \underset{\sim}{1}_j')\underset{\sim}{y}$$

$$= (\underset{\sim}{I}_p \otimes \Sigma^{+}\ \underset{\sim}{1}_j'/n_j)\underset{\sim}{y}$$

$$= (\underset{\sim}{I}_p \otimes \Sigma^{+}\ 1/n_j)(\underset{\sim}{I}_p \otimes \underset{\sim}{1}_j')\underset{\sim}{y}$$

$$= (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y}$$

$$= \hat{\underset{\sim}{\mu}} \quad \text{under} \quad \text{O.L.S.}$$

Note that $\underset{\sim}{\Omega}$ can be arbitrary, although three specific forms of particular interest are as follows:

$$\underset{\sim}{\Omega} = \sigma^2\{(1-\rho)\underset{\sim}{I} + \rho\underset{\sim}{J}\} \tag{3.6}$$

$$\underset{\sim}{\Omega}\ , \quad \text{tri diagonal}\ ; \quad \text{i.e.,}\quad \underset{\sim}{\Omega} = \sigma^2 \begin{bmatrix} 1 & \rho & 0 & \cdots & 0 \\ \rho & 1 & \rho & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{3.7}$$

$$\underset{\sim}{\Omega} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{p-1} \\ \rho & 1 & \rho & \cdots & \rho^{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \cdots & 1 \end{bmatrix} \tag{3.8}$$

The form in (3.6) is the intraclass correlation model, (3.7) assumes that only observations between adjacent points in time are correlated and (3.8) is an example of a Toeplitz matrix [e.g., Press (1972), p. 15].

Estimation of $\underset{\sim}{\mu}$ in (3.1) is easily accomplished since $\underset{\sim}{X}'\underset{\sim}{X}$ is of full rank. Denoting by $\bar{\underset{\sim}{y}}$ the vector of cell means and using (3.2),

$$\hat{\underset{\sim}{\mu}} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y} = \bar{\underset{\sim}{y}}$$

so that

$$\hat{\mu}_{ij} = \bar{y}_{ij\cdot} \ , \quad i=1,\cdots,p, \quad j=1,\cdots,s \ . \tag{3.9}$$

In addition,

$$\begin{aligned} V(\hat{\underset{\sim}{\mu}}) &= (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}' \ \underset{\sim}{\Sigma} \ \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1} \\ &= [\underset{\sim}{I}_p \otimes \Sigma^+ \ 1/n_j][\underset{\sim}{I}_p \otimes (\Sigma^+ \ \underset{\sim}{1}_j')][\underset{\sim}{\Omega}_{pxp} \otimes \Sigma^+ \ \underset{\sim}{I}_{jn}][\underset{\sim}{I}_p \otimes \Sigma^+ \ \underset{\sim}{1}_j][\underset{\sim}{I}_p \otimes \Sigma^+ \ 1/n_j] \\ &= [\underset{\sim}{I}_p \otimes \Sigma^+ \ 1/n_j][\underset{\sim}{\Omega}_{pxp} \otimes \Sigma^+ \ n_j][\underset{\sim}{I}_p \otimes \Sigma^+ \ 1/n_j] \\ &= \underset{\sim}{\Omega}_{pxp} \ \Sigma^+ \ 1/n_j \ , \end{aligned} \tag{3.10}$$

where the Kronecker sums are over $j=1,\cdots,s$.

In introducing the generalized cell means model, some additional properties are to be noted.

3.1. Error Terms

The sum of squares for error (SSE) after fitting (3.1) is

$$SSE = \underset{\sim}{y}'[\underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}']\underset{\sim}{y} \ .$$

Noting that

$$\begin{aligned} \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}' &= [\underset{\sim}{I}_p \otimes \Sigma^+ \ \underset{\sim}{i}_j][\underset{\sim}{I}_p \otimes \Sigma^+ \ 1/n_j][\underset{\sim}{I}_p \otimes \Sigma^+ \ \underset{\sim}{1}_j'] \\ &= [\underset{\sim}{I}_p \otimes \Sigma^+ \ \underset{\sim}{J}_j/n_j] \ , \end{aligned}$$

where $\underset{\sim}{J}_j$ is an $n_j \times n_j$ matrix containing all 1's, it therefore follows that

$$\begin{aligned} SSE &= \underset{\sim}{y}'[\underset{\sim}{I}_{pn} - \underset{\sim}{I}_p \otimes \Sigma^+ \ \underset{\sim}{J}_j/n_j]\underset{\sim}{y} \\ &= \underset{\sim}{y}'[\underset{\sim}{I}_p \otimes \Sigma^+ \ \underset{\sim}{I}_j - \underset{\sim}{I}_p \otimes \Sigma^+ \ \underset{\sim}{J}_j/n_j]\underset{\sim}{y} \\ &= \underset{\sim}{y}'[\underset{\sim}{I}_p \otimes \Sigma^+ \ \underset{\sim}{N}_j]y \end{aligned} \tag{3.11}$$

where

$$\underset{\sim}{N}_j \equiv \underset{\sim}{I}_j - J_j/n_j \tag{3.12}$$

and $\underset{\sim}{I}_j$ is an identity matrix of order n_j.

In defining N_j, note that $\underset{\sim}{N}_j$ is symmetric and idempotent and that $tr(\underset{\sim}{N}_j = r(\underset{\sim}{N}_j) = n_j - 1$.

3.2 Linear combinations of observed cell means are independent of SSE.

For

$$\hat{\underset{\sim}{\mu}} = \bar{\underset{\sim}{y}} \sim N(\underset{\sim}{\mu}, \underset{\sim}{\Omega} \otimes \Sigma^+ \, 1/n_j) \quad ,$$

it follows that for any nun-null vector $\underset{\sim}{k}'$

$$\underset{\sim}{k}'\hat{\underset{\sim}{\mu}} \sim N[\underset{\sim}{k}'\underset{\sim}{\mu}, \underset{\sim}{k}'(\underset{\sim}{\Omega} \otimes \Sigma^+ \, 1/n_j)\underset{\sim}{k}] \quad .$$

Let

$$\underset{\sim}{k}'\hat{\underset{\sim}{\mu}} = \underset{\sim}{k}'[\underset{\sim}{I}_p \otimes \Sigma^+ \, \underset{\sim}{1}_{n_j}/n_j]y = \underset{\sim\sim}{B}\underset{\sim}{y} \quad ,$$

say, and take $\underset{\sim}{A}$ = the matrix of the quadratic form in (3.11) = $[\underset{\sim}{I}_p \otimes \Sigma^+ \, \underset{\sim}{N}_j]$.

Then $\underset{\sim}{y}'\underset{\sim\sim}{A}\underset{\sim}{y}$ and $\underset{\sim\sim}{B}\underset{\sim}{y}$ for $\underset{\sim}{y} \sim N[\underset{\sim}{\mu}, \underset{\sim}{\Sigma}]$ are independent if and only if $\underset{\sim\sim\sim}{B\Sigma A} = 0$ [e.g., Searle, (1971), Theorem 3, p. 59].

Denoting by $\underset{\sim}{1}_j$ the vector $\underset{\sim}{1}$ of order n_j,

$$\underset{\sim\sim\sim}{B\Sigma A} = \underset{\sim}{k}'[\underset{\sim}{I}_p \otimes \Sigma^+ \, \underset{\sim}{1}_j'/n_j][\underset{\sim}{\Omega}_p \otimes \Sigma^+ \, \underset{\sim}{I}_j][\underset{\sim}{I}_p \otimes \Sigma^+ \, \underset{\sim}{N}_j]$$

$$= \underset{\sim}{k}'[\underset{\sim}{\Omega}_p \otimes \Sigma^+ \, \underset{\sim}{1}_j'\underset{\sim}{N}_j/n_j] = \underset{\sim}{0} \quad ,$$

since

$$\underset{\sim}{1}_j'\underset{\sim}{N}_j = \underset{\sim}{1}_j'[\underset{\sim}{I}_j - \underset{\sim}{J}_j/n_j]$$

$$= \underset{\sim}{1}_j' - \underset{\sim}{1}_j' = \underset{\sim}{0}_j \quad ,$$

where $\underset{\sim}{0}_j$ is an $n_j \times n_j$ null matrix.

3.3. Estimable functions and testable hypotheses.

Note that any parametric function of the cell means $\underset{\sim}{q}'\underset{\sim}{\mu}$ will be estimable since it will always be the case that $\underset{\sim}{q}' = \underset{\sim}{t}'\underset{\sim}{X}$ for some $\underset{\sim}{t}'$. In particular, the test for estimability [e.g., Searle, sec. 5.4g, (1971)] will laways be satisfied since $(\underset{\sim}{X}'\underset{\sim}{X})^{-} \equiv (\underset{\sim}{X}'\underset{\sim}{X})^{-1}$ so that $\underset{\sim}{q}' = \underset{\sim}{q}'(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{X} = \underset{\sim}{q}'$. It immediately follows that all hypotheses of the form $H:\underset{\sim}{k}'\underset{\sim}{\mu} = \underset{\sim}{m}$ are testable.

3.4. Estimation under restricted models.

On fitting $\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{\mu} + \underset{\sim}{\epsilon}$ with $V(\underset{\sim}{\epsilon}) = \underset{\sim}{\Omega} \otimes \underset{\sim}{I}$ subject to $\underset{\sim}{P}'\underset{\sim}{\mu} = \underset{\sim}{m}$, the estimation procedure involves minimizing

$$(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\mu})' \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\mu})$$

subject to $\underset{\sim}{P}'\underset{\sim}{\mu} = \underset{\sim}{m}$ on equivalently minimizing

$$(\underset{\sim}{y}-\underset{\sim}{X}\hat{\underset{\sim}{\mu}})' \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{y}-\underset{\sim}{X}\hat{\underset{\sim}{\mu}}) + 2\underset{\sim}{\theta}'(\underset{\sim}{P}'\underset{\sim}{\mu}-\underset{\sim}{m}) \qquad (3.13)$$

where $2\underset{\sim}{\theta}'$ is a vector of hagrange multipliers.

Differentiating (3.13) with respect to $\hat{\underset{\sim}{\mu}}$ and $\underset{\sim}{\theta}$ and setting the results equal to zero leads to the equations

$$\underset{\sim}{X}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{X}\underset{\sim}{\mu} + \underset{\sim}{P}\underset{\sim}{\theta} = \underset{\sim}{X}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{y}$$

$$\underset{\sim}{P}'\underset{\sim}{\mu} = \underset{\sim}{m} \quad .$$

The solution is

$$\hat{\underset{\sim}{\mu}} = \hat{\underset{\sim}{\mu}} - (\underset{\sim}{X}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{X})^{-1}\underset{\sim}{P}[\underset{\sim}{P}'(\underset{\sim}{X}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{X})^{-1}\underset{\sim}{P}]^{-1}[\underset{\sim}{P}'\hat{\underset{\sim}{\mu}} - \underset{\sim}{m}] \quad .$$

For $\underset{\sim}{X} = \underset{\sim}{I}_p \otimes \Sigma^{+} \underset{\sim}{1}_j$ and $\underset{\sim}{\Sigma} = \underset{\sim}{\Omega} \otimes \Sigma^{+} \underset{\sim}{I}_j$ the solution simplifies to

$$\hat{\underset{\sim}{\mu}} = \bar{\underset{\sim}{y}} - [\underset{\sim}{\Omega}^{-1} \otimes \Sigma^{+} \, 1/n_j]\underset{\sim}{P}[\underset{\sim}{P}'(\underset{\sim}{\Omega}^{-1} \otimes \Sigma^{+} \, 1/n_j)\underset{\sim}{P}]^{-1}[\underset{\sim}{P}'\bar{\underset{\sim}{y}} - \underset{\sim}{m}] \; . \qquad (3.14)$$

For the special case

$$\underset{\sim}{\Sigma} = \sigma^2 \underset{\sim}{I}_{psn} = \sigma^2[\underset{\sim}{I}_p \otimes \Sigma^{+} \, \underset{\sim}{I}_j] \; ; \quad \text{i.e.,} \quad \underset{\sim}{\Omega} = \sigma^2 \underset{\sim}{I}_p \; .$$

The solution simplifies further to

$$\hat{\underset{\sim}{\mu}} = \bar{\underset{\sim}{y}} - (\underset{\sim}{I}_p \otimes \Sigma^{+} \, 1/n_j)\underset{\sim}{P}[\underset{\sim}{P}'(\underset{\sim}{I}_p \otimes \Sigma^{+} \, 1/n_j)\underset{\sim}{P}]^{-1}[\underset{\sim}{P}'\underset{\sim}{y} - \underset{\sim}{m}] \; .$$

4. The Two Treatment Two Period Design

On applying (3.1) and (3.3) to the case where p = s = 2, the cell means model is

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk} \tag{4.1}$$

with

$$V(\epsilon) = \underset{\sim}{\Omega} \otimes \underset{\sim}{I}_n \quad ,$$

where

$$\underset{\sim}{\Omega} = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} .$$

From (3.9) and (3.10),

$$\hat{\mu}_{ij} = \bar{y}_{ij.} \quad \text{for} \quad i,j=1,2 \quad ,$$

and

$$V(\hat{\mu}) = \Omega_{2x2} \otimes \sum_{j=1}^{2}{}^{+} 1/n_j \quad . \tag{4.2}$$

4.1. Construction of BLUE's

The μ_{ij}'s in (4.1) differ from one another by the nature of the experimental design. Thus the correspondence between the right-hand sides of (4.1) and (2.2) is such that

$$\mu_{ij} = E(Y_{ijk})$$

where the expectation is taken with respect to model (2.2). Thus

$$\begin{aligned} \mu_{11} &= \tau_1 \\ \mu_{12} &= \tau_2 \\ \mu_{21} &= \tau_2 + \rho_1 \\ \mu_{22} &= \tau_1 + \rho_2 \quad . \end{aligned} \tag{4.3}$$

Note that the μ_{ij}'s are not structurally affected by the $\xi_{k(j)}$'s since, by assumption, $E(\xi_{k(j)}) = 0$. The effect of introducing random effects into a model is to alter the covariance structure of $\underset{\sim}{y}$. The $\xi_{k(j)}$'s are thus taken into account in $\underset{\sim}{\Omega}$ of (4.2). Note also that μ, the effect due to an overall mean, does not appear in (4.3). Since it is a non-estimable effect which is seldom of any real interest, each τ_i in (4.3) has been redefined as $\tau_i + \mu'$ of (2.2).

Since $\hat{\mu}_{ij} = \bar{y}_{ij\cdot}$, the correspondence in (4.3) is such that

$$\begin{aligned} \bar{y}_{11\cdot} &= \hat{\tau}_1 \\ \bar{y}_{12\cdot} &= \hat{\tau}_2 \\ \bar{y}_{21\cdot} &= \hat{\tau}_2 + \hat{\rho}_1 \\ \bar{y}_{22\cdot} &= \hat{\tau}_1 + \hat{\rho}_2 \quad . \end{aligned} \tag{4.4}$$

The BLUE's corresponding to estimable functions in (2.2) are easily obtained by solving (4.4), thereby yielding

$$\widehat{\tau_1 - \tau_2} = \bar{y}_{11\cdot} - \bar{y}_{12\cdot} \tag{4.5}$$

$$\widehat{\rho_1 - \rho_2} = \bar{y}_{11\cdot} + \bar{y}_{21\cdot} - \bar{y}_{12\cdot} - \bar{y}_{22\cdot} \tag{4.6}$$

with

$$V(\widehat{\tau_1 - \tau_2}) = \frac{n\sigma^2}{n_1 n_2} \qquad (\text{recall } n = n_1 + n_2) \tag{4.7}$$

$$V(\widehat{\rho_1 - \rho_2}) = \frac{2n\sigma^2}{n_1 n_2}(1+\rho) \quad . \tag{4.8}$$

Results (4.5)-(4.8) constitute Table 3 in Grizzle (1965). Grizzle's derivation is the less direct but more traditional one: setting up normal

equations for the over-parametrized model, imposing nonestimable constraints, and solving. There is clearly a computational advantage in the cell means approach.

4.2. Sums of Squares for Main Effects

The sums of squares corresponding to testing H: $\tau_1 - \tau_2 = 0$ and H: $\rho_1 - \rho_2 = 0$ can be constructed in two ways.

Since $\widehat{\tau_1 - \tau_2}$ and $\widehat{\rho_1 - \rho_2}$ are one degree of freedom contrasts of means, it follows immediately [e.g., Snedecor and Cochran (1967), Chap. 12] that

$$\text{SSD} = \text{SS(direct effects)} = [\Sigma\, 1/n_j]^{-1}[\bar{y}_{11.} - \bar{y}_{12.}]^2$$

$$= (n_1 n_2/n)[\bar{y}_{11.} - \bar{y}_{12.}]^2 \qquad (4.9)$$

and

$$\text{SSR} = \text{SS(residual effects)} = [\Sigma\, 2/n_j]^{-1}[\bar{y}_{11.} + \bar{y}_{21.} - \bar{y}_{12.} - \bar{y}_{22.}]^2$$

$$= (n_1 n_2/2n)[\bar{y}_{11.} + \bar{y}_{21.} - \bar{y}_{12.} - \bar{y}_{22.}]^2. \qquad (4.10)$$

The corresponding expectations are then

$$E(\text{SSD}) = E(\text{MSD}) = (n_1 n_2/n)[\mu_{11} - \mu_{12}]^2 + \sigma^2 \qquad (4.11)$$

$$E(\text{SSR}) = E(\text{MSR}) = (n_1 n_2/2n)[\mu_{.1} - \mu_{.2}]^2 + \sigma^2\,(1+\rho)\,. \qquad (4.12)$$

A more general derivation appeals to the theory for testing a general linear hypothesis [e.g., Searle (1971), Chap. 7, Sec. 5]. Thus

$$\text{H: } \tau_1 - \tau_2 = 0$$

is equivalent to

$$\text{H: } \underset{\sim}{k}'\underset{\sim}{\mu} = 0 \qquad (4.13)$$

for

$$\underset{\sim}{k}' = (1 \quad -1 \quad 0 \quad 0) \quad .$$

Since

$$\hat{\underset{\sim}{\mu}} \equiv \bar{\underset{\sim}{y}} \sim N(\underset{\sim}{\mu} \; \underset{\sim}{\Omega} \otimes \Sigma^{+} \; 1/n_j)$$

it follows that

$$\underset{\sim}{k}'\hat{\underset{\sim}{\mu}} \sim N[\underset{\sim}{k}'\underset{\sim}{\mu}, \; \underset{\sim}{k}'(\underset{\sim}{\Omega} \otimes \Sigma^{+} \; 1/n_j)\underset{\sim}{k}] \; ,$$

and since

$$\underset{\sim}{k}'(\underset{\sim}{\Omega} \otimes \Sigma^{+} \; 1/n_j)\underset{\sim}{k} = n\sigma^2/n_1 n_2 \quad ,$$

$$\underset{\sim}{k}'\hat{\underset{\sim}{\mu}} \sim N[\underset{\sim}{k}'\underset{\sim}{\mu}, \; n\sigma^2/n_1 n_2] \quad .$$

Taking $n_1 n_2/n$ as the 1×1 matrix of the quadratic form corresponding to (4.13),

$$SSD = (n_1 n_2/n)\bar{\underset{\sim}{y}}'[\underset{\sim\sim}{k}\underset{\sim}{k}']\bar{\underset{\sim}{y}} = (n_1 n_2/n)[\bar{y}_{11.} - \bar{y}_{12.}]^2 \quad . \tag{4.14}$$

Note that SSD in (4.14) is

$$\bar{\underset{\sim}{y}}'(n_1 n_2/n) \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \bar{\underset{\sim}{y}} = \bar{\underset{\sim}{y}}'\underset{\sim}{A}\bar{\underset{\sim}{y}} \quad , \quad \text{say.}$$

The quadratic form is distributed proportional to a $\chi^2(1)$ since

$$[\underset{\sim}{A} \; \text{var}(\underset{\sim}{y})]/\sigma^2 = \frac{n_1 n_2}{n} \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/n_1 & 0 & \rho/n_1 & 0 \\ 0 & 1/n_2 & 0 & \rho/n_2 \\ \rho/n_1 & 0 & 1/n_1 & 0 \\ 0 & \rho/n_2 & 0 & 1/n_2 \end{bmatrix}$$

which is an idempotent matrix.

Finally, note that $tr[A\ var(\bar{\underset{\sim}{y}})] = \sigma^2$ and $\mu'A\mu = (n_1n_2/n)[\mu_{11} - \mu_{12}]^2$. Thus

$$SSD = (n_1n_2/n)[\bar{y}_{11.} - \bar{y}_{12.}]^2 \sim \sigma^2\chi^{2\prime}[1,(n_1n_2/n)(\mu_{11} - \mu_{12})^2] \quad .$$

In a similar fashion, it can be verified that for SSR of (4.10)

$$SSR \sim \sigma^2(1+\rho)\chi^{2\prime}[1,(n_1n_2/2n)(\mu_{.1} - \mu_{.2})^2] \quad .$$

4.3. The Error Term

A key difficulty in the analysis of mixed models lies in choosing the proper error terms for testing hypotheses. Many experimenters, not blessed with a Fisherian sense of what goes with what, often feel uneasy about constructing sums of squares for errors. In addition to providing an exceptionally easy procedure for obtaining BLUE's of estimable functions of parameters in the over-parametrized model, the cell means formulation enables one to easily unravel components of the error term.

Recall from (3.11) that

$$SSE = \underset{\sim}{y}'[\underset{\sim}{I}_p \otimes \Sigma^+ \underset{\sim}{N}_j]\underset{\sim}{y}$$

and, for the 2 x 2 case,

$$SSE = \underset{\sim}{y}'[\underset{\sim}{I}_2 \otimes \sum_{j=1}^{2}{}^{+} \underset{\sim}{N}_j]\underset{\sim}{y} \tag{4.15}$$

$$= \underset{\sim}{y}'[\underset{\sim}{I}_2 \otimes (\underset{\sim}{N}_1 \otimes \underset{\sim}{N}_2)]\underset{\sim}{y} \tag{4.16}$$

$$= \underset{\sim}{y}' \begin{bmatrix} \underset{\sim}{N}_1 & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{N}_2 & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{N}_1 & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{N}_2 \end{bmatrix} \underset{\sim}{y} \quad ,$$

and where, as in (3.12),

$$\underset{\sim}{N}_j = \underset{\sim}{I}_j - (1/n_j)\underset{\sim}{J}_j \quad , \quad j=1,2 \quad .$$

SSE in (4.15) is the pooled within cells error term; i.e.,

$$SSE = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{n_j} (Y_{ijk} - \bar{Y}_{ij.})^2 \quad .$$

As an aid in establishing distributional properties of the components of (4.15) some additional notation is necessary. Let $\underset{\sim}{Z}_i$ denote a square matrix whose only nonzero element is a 1 in the i,i^{th} position. For example, for order 2,

$$\underset{\sim}{Z}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \underset{\sim}{Z}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad .$$

Note that for arbitrary n,

$$\underset{\sim}{Z}_i\underset{\sim}{Z}_j = \begin{cases} \underset{\sim}{Z}_i \ , & i=j \\ \underset{\sim}{0}_n \ , & i \neq j \end{cases} \tag{4.17}$$

Thus from (4.15)

$$SSE = \underset{\sim}{y}'[\underset{\sim}{I}_2 \otimes \Sigma^+ \underset{\sim}{N}_j]\underset{\sim}{y}$$

$$= \underset{\sim}{y}'[(\underset{\sim}{Z}_1 + \underset{\sim}{Z}_2) \otimes \Sigma^+ \underset{\sim}{N}_j]\underset{\sim}{y}$$

which can be rewritten as

$$SSE = \sum_{i=1}^{2} \sum_{j=1}^{2} [\underset{\sim}{y}'(\underset{\sim}{Z}_i \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_k)\underset{\sim}{y}] , \tag{4.18}$$

where δ_{kj} is the usual Kronecker delta. Thus

$$\begin{aligned} SSE(ij) &= SSE(\text{within cell } i,j) \\ &= \underset{\sim}{y}'(\underset{\sim}{Z}_i \otimes \Sigma^{+} \delta_{kj} \underset{\sim}{N}_j)\underset{\sim}{y} \\ &= \underset{\sim}{y}' \underset{\sim}{A}_{ij} \underset{\sim}{y} \quad . \end{aligned} \tag{4.19}$$

Since $\underset{\sim}{Z}_i$ and $\underset{\sim}{N}_k$ are symmetric and idempotent and since $\underset{\sim}{A}_{ij}$ is block diagonal, it immediately follows that $\underset{\sim}{A}_{ij}$ is symmetric and idempotent.

Consider

$$\begin{aligned} \underset{\sim}{A}_{ij} \underset{\sim}{\Sigma} &= [\underset{\sim}{Z}_i \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_k][\underset{\sim}{\Omega} \otimes \Sigma^{+} \underset{\sim}{I}_j] \\ &= [\underset{\sim}{Z}_i \underset{\sim}{\Omega} \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_j] \quad . \end{aligned}$$

Idempotency of $(1/\sigma^2)\underset{\sim}{A}_{ij} \underset{\sim}{\Sigma}$ follows from noting that

$$\underset{\sim}{A}_{ij} = \underset{\sim}{Z}_i \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_k$$

$$\underset{\sim}{A}_{ij} \underset{\sim}{\Sigma} = \{\underset{\sim}{Z}_i \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_k\}\{\underset{\sim}{\Omega} \otimes \sum_{k=1}^{2}{}^{+} \underset{\sim}{I}_k\}$$

$$= \underset{\sim}{Z}_i \underset{\sim}{\Omega} \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_k$$

$$(\underset{\sim}{A}_{ij} \underset{\sim}{\Sigma})^2 = \underset{\sim}{Z}_i \underset{\sim}{\Omega})^2 \otimes \left(\sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_k \right)^2$$

$$= \underset{\sim}{Z}_i \underset{\sim}{\Omega} \underset{\sim}{Z}_i \underset{\sim}{\Omega} \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj}^2 \underset{\sim}{N}_k^2$$

$$= \underset{\sim}{Z}_i \underset{\sim}{\Omega} \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_j \quad ,$$

since $\delta_{kj}^2 = \delta_{kj}$, $\underset{\sim}{N}_k^2 = \underset{\sim}{N}_k$ and $\underset{\sim}{Z}_i \underset{\sim}{\Omega} \underset{\sim}{Z}_i \underset{\sim}{\Omega} = \underset{\sim}{Z}_i \underset{\sim}{\Omega}$ by virtue of $\underset{\sim}{\Omega}$; e.g.,

$$\underset{\sim}{Z}_1 \underset{\sim}{\Omega} = \sigma^2 \begin{pmatrix} 1 & \rho \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \underset{\sim}{Z}_2 \underset{\sim}{\Omega} = \sigma^2 \begin{pmatrix} 0 & 0 \\ \rho & 1 \end{pmatrix}$$

so that $(1/\sigma^2)\underset{\sim}{Z}_1 \underset{\sim}{\Omega}$ and $(1/\sigma^2)\underset{\sim}{Z}_2 \underset{\sim}{\Omega}$ are both idempotent.

Finally, since $\operatorname{tr}(\underset{\sim}{A}_{ij} \underset{\sim}{\Sigma}) = (n_j - 1)\sigma^2$, it follows from the above that

$$\underset{\sim}{y}' \underset{\sim}{A}_{ij} \underset{\sim}{y} \sim \sigma^2 \chi^2 (n_j - 1) \ .$$

The individual quadratic forms are not distributed independently. Consider the two quadratic forms within sequence j; i.e.,

$$\underset{\sim}{A}_{ij} = \underset{\sim}{Z}_i \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_k \qquad \text{for} \quad i=1,2 \ .$$

$$\underset{\sim}{A}_{1j} \underset{\sim}{\Sigma} \underset{\sim}{A}_{2j} = [\underset{\sim}{Z}_1 \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_k][\underset{\sim}{\Omega} \otimes \sum_{k=1}^{2}{}^{+} \underset{\sim}{I}_k][\underset{\sim}{Z}_2 \otimes \sum_{k=1}^{2}{}^{+} \delta_{kj} \underset{\sim}{N}_k]$$

$$= \underset{\sim}{Z}_1 \underset{\sim}{\Omega} \underset{\sim}{Z}_2 \otimes \Sigma^{+} \delta_{kj} \underset{\sim}{N}_k \neq 0 \quad ,$$

since

$$\underset{\sim}{Z}_1 \underset{\sim}{\Omega} \underset{\sim}{Z}_2 = \sigma^2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \sigma^2 \begin{pmatrix} 0 & 0 \\ 0 & \rho \end{pmatrix} \ .$$

Quadratic forms between sequences are independent since, for $j \neq j'$,

$$\underset{\sim}{A}_{ij}\,\underset{\sim}{\Sigma}\,\underset{\sim}{A}_{i'j'} = [\underset{\sim}{Z}_i \otimes \sum_{k=1}^{2}\Sigma^{+}\delta_{kj}\underset{\sim}{N}_k][\underset{\sim}{\Omega} \otimes \sum_{k=1}^{2}\Sigma^{+}\underset{\sim}{I}_k][\underset{\sim}{Z}_{i'} \otimes \sum_{k=1}^{2}\Sigma^{+}\delta_{kj}\underset{\sim}{N}_k] = \underset{\sim}{0} \quad ,$$

by virtue of

$$\left(\sum_{k=1}^{2}\Sigma^{+}\,\delta_{kj}\underset{\sim}{N}_k\right)\left(\sum_{k=1}^{2}\Sigma^{+}\,\delta_{kj'}\underset{\sim}{N}_k\right) = \underset{\sim}{0} \quad .$$

Finally note that SSE in (4.15) has expectation $(2n-4)\sigma^2$ but it is not distributed proportional to a central chi square. This latter fact is seen immediately by noting that the product of the matrix of the quadratic forms in (4.15) and $\underset{\sim}{\Sigma}$ in (3.3) is $(\Omega \otimes \sum_{k=1}^{2}\Sigma^{+}\,\underset{\sim}{N}_k)$ which is not idempotent.

The above results are assembled into two summary tables below.

Table 4.1

Partition of Error Term: Within Periods

Source	d.f.	SS*
Error	2n-4	SSE
Within Period 1	n-2	SSP1
Within Sequence 1	n_1-1	SS(11)
Within Sequence 2	n_2-1	SS(12)
Within Period 2	n-2	SSP2
Within Sequence 1	n_1-1	SS(21)
Within Sequence 2	n_2-1	SS(22)

* Each corresponding mean square has expectation σ^2. All sums of squares are distributed proportional to central χ^2's with the exception of SSE. Sums of squares such as SS(ij) are distributed independently only if they are from different sequences, e.g., SS(ij) and SS(i'j') are distributed independently for $j \neq j'$.

Table 4.2

Partition of Error Term: Within Sequences

Source	d.f.	SS*	Distribution
Error	$2n-4$	SSE	not χ^2
Within Sequence 1	$2n_1-2$	SS1	not χ^2
Within Period 1	n_1-1	SS(11)	χ^2
Within Period 2	n_1-1	SS(21)	χ^2
Within Sequence 2	$2n_2-2$	SS2	not χ^2
Within Period 1	n_2-1	SS(12)	χ^2
Within Period 2	n_2-1	SS(22)	χ^2

* Each corresponding mean square has expectation σ^2.

The following pairs of quadratic forms are distributed independently: SS1 and SS2, SS(i,j) and SS(i'j') for $j \neq j'$, where the SS(ij) sums of squares are the same as in Table 4.1.

From either of Tables 4.1 or 4.2, tests of homogeneity of the within cell error terms can be formulated by considering ratios of independent within cell mean squares. For example, if the model were such that $\text{var}(y_{ijk}) = \sigma^2_{ij}$ then one might wish to test

$$H: \sigma^2_{ij} = \sigma^2 \quad \text{for all } i,j \ .$$

This can be accomplished by noting that

$$F_1 = MS(11)/MS(12) \quad \text{tests} \quad H: \sigma^2_{11} = \sigma^2_{12} \ ,$$

$$F_2 = MS(11)/MS(22) \quad \text{tests} \quad H: \sigma^2_{11} = \sigma^2_{22} \ ,$$

$$F_3 = MS(21)/MS(12) \quad \text{tests} \quad H: \sigma^2_{21} = \sigma^2_{12} \ ,$$

and

$$F_4 = MS(21)/MS(22) \quad \text{tests} \quad H: \sigma^2_{21} = \sigma^2_{22} \ .$$

Note that in order to test H: $\sigma^2_{ij} = \sigma^2$ for all i,j it is necessary to use any three of F_1 - F_4. A simultaneous error rate of α can be achieved by making each individual test at the $\alpha' = 2/3$ level.

An alternative (and more useful) partitioning of SSE is obtained as follows. Define $\underset{\sim}{K}_2 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = 2\underset{\sim}{I}_2 - \underset{\sim}{J}_2$, and note that, similar to J matrices, $\underset{\sim}{K}_2^2 = 2\underset{\sim}{K}$, $(\frac{1}{2}\underset{\sim}{K}_2)^2 = \frac{1}{2}\underset{\sim}{K}$. Observe that

$$\underset{\sim}{I}_2 = \tfrac{1}{2}\underset{\sim}{J}_2 + \tfrac{1}{2}\underset{\sim}{K}_2 \ .$$

Thus

$$\begin{aligned} SSE &= \underset{\sim}{y}'(\underset{\sim}{I}_2 \otimes \Sigma^+ \underset{\sim}{N}_k)y \\ &= \underset{\sim}{y}'(\tfrac{1}{2}\underset{\sim}{J}_2 \otimes \Sigma^+ \underset{\sim}{N}_k)\underset{\sim}{y} + \underset{\sim}{y}'(\tfrac{1}{2}\underset{\sim}{K}_2 \otimes \Sigma^+ \underset{\sim}{N}_k) \qquad (4.20) \\ &= SS(\text{sub:seq}) + SS(\text{period} \times \text{sub:seq}) \\ &= SS(S) \qquad + SS(P \times S) \ . \end{aligned}$$

The right-hand side of (4.20) can be partitioned furhter by using the Kronecker delta operators as before; thus

$$\begin{aligned} SSE = &\sum_{j=1}^{2} \underset{\sim}{y}'(\tfrac{1}{2}\underset{\sim}{J}_2 \otimes \sum_{k=1}^{2} \Sigma^+ \delta_{jk}\underset{\sim}{N}_k)\underset{\sim}{y} \\ &+ \sum_{j=1}^{2} \underset{\sim}{y}'(\tfrac{1}{2}\underset{\sim}{K}_2 \otimes \sum_{k=1}^{2} \Sigma^+ \delta_{jk}\underset{\sim}{N}_k)y \qquad (4.21) \\ = SS(S) &\qquad + SS(PS) \end{aligned}$$

$$= SS(\text{sub:seq 1}) + SS(\text{sub:seq 2})$$

$$+ SS(P \times \text{sub:seq 1}) + SS(P \times \text{sub:seq 2})$$

$$= SS(S1) + SS(S2) + SS(PS1) + SS(PS2) .$$

On noting that $\underset{\sim}{K}_2\underset{\sim}{J}_2 = \underset{\sim}{0}$, $\underset{\sim}{J}_2\underset{\sim}{\Omega} = \sigma^2(1+\rho)\underset{\sim}{I}_2$, and $\underset{\sim}{K}_2\underset{\sim}{\Omega} = \sigma^2(1-\rho)\underset{\sim}{I}_2$, it easily follows that the two quadratic forms in (4.20) are independently distributed as proportional to χ^2.

Similarly, the quadratic forms in (4.21) are all distributed proportional to χ^2'2, and SS(S1) is independent of SS(S2) and SS(PS1) is independent of SS(PS2).

Using the assumptions of independence of subjects between sequences, the following table results.

Table 4.3

Partition of Error Term: Subjects Within Sequences

Source	d.f.	SS	E(MS)
SSE	2n-r	SSE	σ^2
Subjects wn Sequences	n-2	SS(S)	$\sigma^2(1+\rho)$
Subjects:Seq 1	n_1-1	SS1	$\sigma^2(1+\rho)$
Subjects:Seq 2	n_2-1	SS2	$\sigma^2(1+\rho)$
Period X Subjects:Sequences	n-2	SS(PS)	$\sigma^2(1-\rho)$
Period X Subjects:Seq 1	n_1-1	SSP1	$\sigma^2(1-\rho)$
Period X Subjects:Seq 2	n_2-1	SSP2	$\sigma^2(1-\rho)$

Excluding SSE all of the above sums of squares are independently distributed as proportional to central χ^2's.

An additional hypothesis that is testable is that subjects within sequences share a common variance. An F-statistic for doing so is MSS1/MSS2 ~ $F(n_1-1,n_2-1)$ or equivalently MSSP1/MSSP2 ~ $P(n_1-1,n_2-1)$. Failure to reject the hypothesis of common variances within sequences provides justification for pooling the individual sums of squares into SS(S) and SS(PS). Since these sums of squares are also independently distributed as proportional to central χ^2's, the central F statistic formed by MSS(S)/MSS(PS) provides a test of H: $(1+\rho)(1-\rho) = 0$, which is equivalent to H: $\rho = 0$, which in turn is equivalent to H: $\sigma_s^2 = 0$.

Unbiased estimators of σ_e^2, σ_s^2, and ρ can be obtained by equating the pooled mean squares MSS(S) and MS(PS) as well as MSE to their expectations; e.g.,

$$\text{MSE} = \sigma^2 \ ,$$

$$\text{MSS(S)} = \sigma^2(1+,) = \sigma^2 + \sigma_s^2 \ ,$$

$$\text{MS(PS)} = \sigma^2(1-\rho) = \sigma^2 - \sigma_s^2 \ .$$

Some unbiased estimators are therefore

$$\hat{\rho} = \frac{\text{MSS(S)}}{\text{MSE}} - 1 = 1 - \frac{\text{MSS(PS)}}{\text{MSE}}$$

$$\hat{\sigma}^2 = \text{MSE}$$

$$\hat{\sigma}_s^2 = \text{MSS(S)} - \text{MSE} = \text{MSE} - \text{MSS(PS)}$$

and

$$\hat{\sigma}_e^2 = 2\text{MSE} - \text{MSS(S)} = 2\text{MSS(PS)} - \text{MSE} \ .$$

4.4. The No Residual Effects Model

Under the assumption of no residual effects, the estimable functions in (2.1) are $\tau_1 - \tau_2$, $\pi_1 - \pi_2$ and $\delta_1 - \delta_2$. The cell means model is still (4.1);

However, the structure of the μ_{ij}'s is now such that

$$\mu_{11} = \tau_1 + \pi_1 + \delta_1$$

$$\mu_{12} = \tau_2 + \pi_1 + \delta_2$$

$$\mu_{21} = \tau_2 + \pi_2 + \delta_1$$

$$\mu_{22} = \tau_1 + \pi_2 + \delta_2 \quad . \tag{4.22}$$

Using $\hat{\mu}_{ij} = \bar{y}_{ij.}$, equations (4.22) lead to the following. BLUE's of estimable functions:

$$\tau_1 - \tau_2 = \tfrac{1}{2}(\bar{y}_{11.} + \bar{y}_{22.} - \bar{y}_{12.} - \bar{y}_{21.})$$

$$\pi_1 - \pi_2 = \tfrac{1}{2}(\bar{y}_{11.} + \bar{y}_{12.} - \bar{y}_{21.} - \bar{y}_{22.}) \quad ,$$

$$\delta_1 - \delta_2 = \tfrac{1}{2}(\bar{y}_{11.} + \bar{y}_{21.} - \bar{y}_{12.} - \bar{y}_{22.}) \quad .$$

The corresponding sums of squares are

$$\text{SSD} = \text{SS(dis.eff.)} = (2n_1n_2)/n\ (\bar{y}_{11.} + \bar{y}_{22.} - \bar{y}_{12.} - \bar{y}_{21.})^2 \quad , \tag{4.23}$$

$$\text{SSP} = \text{SS(periods)} = (2n_1n_2)/n\ (\bar{y}_{11.} + \bar{y}_{12.} - \bar{y}_{21.} - \bar{y}_{22.})^2 \quad , \tag{4.24}$$

and

$$\text{SSS} = \text{SS(sequences)} = (2n_1n_2)/n\ (\bar{y}_{11.} + \bar{y}_{21.} - \bar{y}_{12.} - \bar{y}_{22.})^2 \quad , \tag{4.25}$$

having expectations

$$E[\text{SS/dir.eff.})] = (2n_1n_2)/n\ (\mu_{11} + \mu_{22} - \mu_{12} - \mu_{21})^2 + \sigma^2(1-\rho) \quad , \tag{4.26}$$

$$E[\text{SS(periods)}] = (2n_1n_2)/n\ (\mu_{1.} - \mu_{2.})^2 + \sigma^2(1-\rho) \quad , \tag{4.27}$$

and

$$E[\text{SS(sequences)} = (2n_1n_2)/n\ (\mu_{.1} - \mu_{.2})^2 + \sigma^2(1+\rho) \quad , \tag{4.28}$$

respectively.

Collecting (4.23)-(4.28) together with the error terms in Table 4.3 gives the analysis of variance below.

Table 4.4

Analysis of Variance for Model with no Residual Effects

Source	d.f.	SS	E(MS)
Direct effects	1	SSD	$\frac{2n_1n_2}{n}(\mu_{11}-\mu_{12}-\mu_{21}+\mu_{22})^2 + \sigma^2(1-\rho)$
Periods	1	SSP	$\frac{2n_1n_2}{n}(\mu_{1.}-\mu_{2.})^2 + \sigma^2(1-\rho)$
Sequences	1	SSS	$\frac{2n_1n_2}{n}(\mu_{.1}-\mu_{.2})^2 + \sigma^2(1+\rho)$
Subjects: Sequences	n-2	SS(S)	$\sigma^2(1+\rho)$
Period X Subjects: Sequences	n-2	SS(PXS)	$\sigma^2(1-\rho)$

Grizzle's (1965) analysis of variance table (Table 5 in his paper) appears to have some errors. The term labeled "Subjects" in Grizzle's (1965) table corresponds to pooling the sums of squares for sequences and subjects: sequences in the above table. Grizzle's (1965) error term is the period X subjects: sequences interaction sum of squares. If the three main effects sums of squares are tested against "error" in Grizzle's (1965) table, only two ratios, namely those corresponding to treatments and periods will in fact be a central F statistic under the appropriate null hypothesis. The F-test corresponding to "subjects" will, strictly speaking, be incorrect, since it will not be testing H: $\sigma_s^2 = 0$ but rather H: $(\sigma_s^2$ + some function of $\delta_1 - \delta_2) = 0$. Only under H: $\delta_1 = \delta_2$

will the F-test for subjects be correct in Grizzle's analysis.

Schematically, the correspondence between Grizzle's (1965) table 5 and Table 4.4 above is as follows.

Table 4.5

Analysis of Variance
(Residual Effects Omitted)

Grizzle Table 5				Table 4.4			
Source	d.f.	F-test	Hypothesis tested	Source	d.f.	F-test	Hypothesis tested
Direct effects	1	←	$\tau_1 = \tau_2$	Direct effects	1	←	$\tau_1 = \tau_2$
Periods	1	←	$\pi_1 = \pi_2$	Periods	1	←	$\pi_1 = \pi_2$
Subjects	n-1	←	?	Sequences	1	←	$\delta_1 = \delta_2$
				Sub:Sequences	n-2		
Error	n-2*	←		Error = Period X Sub: Sequences	n-2	←	

* Degrees of freedom in Grizzle's (1965) table are incorrect.

It is important to note the part that ρ plays in determining the efficiency of the corssover design. Gains in efficiency with the crossover design occur only when $\rho > 0$ for then $E[MS(\text{period X sub:seq})] < \sigma^2$ (see Table 4.3). However, if $\rho < 0$ then the error term for testing direct effects is larger than σ^2 and hence the crossover is less efficient than a comparable completely randomized design. When $\rho = 0$ then MS(sub:seq) and MS(period X sub:seq) are both estimating σ^2 and it would be better to pool these error terms to give one with 2n-4 degrees of freedom. In this case the crossover and completely randomized designs have the same efficiency.

5. A Note on Grizzle's Example

In giving an example to illustrate the analysis of variance procedures Grizzle (1965, sec. 5) considers a set of data consisting of differences between pre-treatment and post-treatment hemoglobin values. His analysis (under a residual effects model) produces a negative estimate for ρ, thus tending to indicate that a completely randomized design would have been more efficient.

It is important to realize that the correlation structure of the observations is sensitive to linear transformations such as taking differences between pre- and post-treatment values. That such a transformation under any of the correlation structures (3.6)-(3.8) (see page 11) will give $\rho \leq 0$ is easily established.

Let the pre-treatment values at periods 1 and 2 be denoted by Y_{1jk} and Y_{3jk}, respectively, and the post-treatment values by Y_{2jk} and Y_{4jk}.

On defining

$$Z_{ijk} = Y_{ijk} - Y_{i'jk} \ , \quad \text{where } i=2,4 \text{ and } i' = i-1 \ ,$$

it follows that

$$\operatorname{cov}(Z_{2jk}, Z_{4jk}) = \operatorname{cov}(Y_{2jk} - Y_{1jk}, Y_{4jk} - Y_{3jk})$$

$$= \begin{cases} 0 & , \text{ under (3.6)} \ , \\ -\rho\sigma^2 & , \text{ under (3.7)} \ , \\ -\rho\sigma^2(\rho-1)^2 & , \text{ under (3.8)} \ . \end{cases}$$

6. References

Brown, B. W. (1978). Statistical controversies in the design of clinical trials. Tech. Report No. 37. Division of Biostatistics, Stanford University, Stanford, California.

Graybill, F. A. (1976). Theory and Application of the Linear Model. Duxbury Press, North Scituate, Massachusetts.

Grizzle, J. E. (1965). The two period change-over design and its use in clinical trials. Biometrics 21:467-480. (Corrigenda: Biometrics 30:727.)

Hocking, R. R. and Speed, F. M. (1975). A full rank analysis of some linear model problems. JASA 70:706-712.

Press, S. J. (1972). Applied Multivariate Analysis. Holt, Rinehart and Winston, Inc., New York.

Searle, S. R. (1966). Matrix Algebra for the Biological Sciences. John Wiley and Sons, Inc., New York.

Searle, S. R. (1971). Linear Models. John Wiley and Sons, Inc., New York.

Snedecor, G. W. and Cochran, W. G. (1967). Statistical Methods. The Iowa State University Press, Ames, Iowa.

Speed, F. M. and Hocking, R. R. (1976). The use of the R( ) notation with unbalanced data. The Amer. Statist. 30:30-34.

Speed, F. M., Hocking, R. R. and Hackney, O. P. (1978). Methods of analysis of linear models with unbalanced data. JASA 73:105-112.